1

Multi-time scale analysis of the spatial representativeness of *in situ* soil moisture data within satellite footprints

B. Molero[1], D. J. Leroux[2], P. Richaume[1], Y. H. Kerr[1], O. Merlin[1], M. H. Cosh[3], R. Bindlish[4]

[1]CESBIO (Centre d'Études Spatiales de la BIOsphère), Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France
[2] CNRM (Centre National de la Recherche Météorologique), UMR 3589 (Météo-France, CNRS), Toulouse, France
[3]USDA ARS Hydrology and Remote Sensing Laboratory, USA
[4]NASA Goddard Space Flight Center, Greenbelt, MD

Corresponding author: Beatriz Molero (beatriz.molero@cesbio.cnes.fr)

Key Points:

A) The spatial representativeness of *in situ* stations tends to increase with the time scale within the

satellite footprint

B) Stations poorly represent the satellite footprint at sub-weekly scales, while either very well or

very poorly at seasonal scales

C) The wavelet correlation (WCor) method is a useful tool to study the spatial scale mismatch

between *in situ* and satellite observations

Key Words: soil moisture, spatial representativeness, time scales, spatial scales, wavelet decomposition,

satellite validation

Abstract

We conduct a novel comprehensive investigation that seeks to prove the connection between spatial and

time scales in surface soil moisture (SM) within the satellite footprint (~50 km). Modeled and measured

point series at Yanco and Little Washita *in situ* networks are first decomposed into anomalies at time

scales ranging from 0.5 to 128 days, using wavelet transforms. Then, their degree of spatial

representativeness is evaluated on a per time-scale basis by comparison to large-spatial scale datasets

(the *in situ* spatial average, SMOS, AMSR2 and ECMWF). Four methods are used for this: temporal

stability analysis (TStab), triple collocation (TC), the percentage of correlated areas (CArea) and a new

proposed approach that uses wavelet-based correlations (WCor). We found that the mean of the spatial

representativeness values tends to increase with the time scale but so does their dispersion. Locations

exhibit poor spatial representativeness at scales below 4 days, while either very good or poor

representativeness at seasonal scales. Regarding the methods, TStab cannot be applied to the anomaly

series due to their multiple zero-crossings and TC is suitable for week and month scales but not for other

scales where datasets cross-correlations are found low. In contrast, WCor and CArea give consistent

results at all time-scales. WCor is less sensitive to the spatial sampling density, so it is a robust method

that can be applied to sparse networks (1 station per footprint). These results are promising to improve

the validation and downscaling of satellite SM series and the optimization of SM networks.

1   Introduction

Soil moisture (SM) plays an important role in atmospheric, hydrologic and ecological processes

[*Rodriguez-Iturbe*, 2000; *Daly and Porporato*, 2005; *Legates et al.*, 2011]. By means of them, it

participates at various scales, from the largest climatic and meteorological scales [*Douville*, 2004;

44   *Drusch*, 2007] to the medium hydrological scale [*Chen et al.*, 2011; *Draper et al.*, 2012] and the

45   smallest field and local scales [*Vereecken et al.*, 2014].

46   The spatial scale of a set of spatially distributed SM measurements (or observations) refers to a triplet of

47   entities: the extent (the area enclosing all the measurements), the spacing (the distance between

48   measurements) and the support (the area actually sensed by the sensor) [*Blöschl and Sivapalan*, 1995].

49   A typical *in situ* station has a support of just some few centimeters (*point* or *local* support). In practice, it

50   represents larger areas because the factors driving SM variability (vegetation, soil texture, topography,

51   rainfall) are spatially connected. This *effective support* or *spatial representativeness area* is defined by

52   the surrounding area showing sufficient similarity with the station location in terms of SM, according to

53   a given evaluation methodology. Hereafter, we will use simply *representativeness* to refer to *spatial*

54   *representativeness*. From space, passive microwave sensors provide SM estimates at a global extent

55   with a resolution (support) of several tens of km, which is defined by the antenna footprint as the area

56   containing half of the total signal power. C- and X-band sensors like AMSR-E, AMSR2 and WindSat

57   [*Wagner et al.*, 2007; *Mladenova et al.*, 2011; *Parinussa et al.*, 2012] and L-band sensors like SMOS

58   and SMAP [*Al Bitar et al.*, 2012; *Kerr et al.*, 2016; *Colliander et al.*, 2017] have shown good skills in

59   capturing the temporal patterns of top-surface SM at ~1 cm and ~5 cm depth, respectively.

60   Factors driving SM variability (vegetation, soil texture, topography, rainfall), although spatially

61   dependent, are not homogeneous within satellite footprints. As a consequence, ground stations rarely

62   represent satellite footprints perfectly. This *spatial scale mismatch* is by principle not known and

63   difficult to estimate. Validation of satellite products usually consists in their direct comparison with *in*

64   *situ* time series through linear metrics (correlation, bias, RMSE). Since the spatial scale mismatch is not

65   considered, the statistics can be hampered to a great extent [*Loew and Schlenz*, 2011; *Crow et al.*, 2012].

66 The satellite-*in situ* spatial scale mismatch can be reduced with upscaling approaches that increase the

67 effective *in situ* support. They can be applied if multiple *in situ* stations are available within the footprint

68 (dense networks). The simplest techniques consist of linear and weighted spatial averages of the stations

69 time series [*Jackson et al.*, 2010]. Stations locations can also be selected in a spatial configuration that

70 ensures the representativeness of the average, based on prior knowledge on, for example, soil texture

71 and land cover [*Bircher et al.*, 2012]. Downscaling of satellite observations can potentially help

72 reducing the spatial scale mismatch for satellite validation [*Malbéteau et al.*, 2016]. The principal

73 drawback of most upscaling and downscaling approaches is the difficulty to assess the method

74 uncertainty and the remaining spatial scale mismatch. When the statistical spatial structure of SM can be

75 inferred, the upscaling uncertainty can be estimated with geostatistical techniques like block kriging

76 [*Wang et al.*, 2015]. However, they need dense sampling schemes (>100, [*Webster & Oliver*, 1992]) that

77 could never be met in practice for long-term *in situ* networks.

78 An alternative approach is to choose directly the ground station that behaves most like the footprint time

79 series. Temporal stability analysis [*Vachaud et al.*, 1985] selects the station that exhibits the smallest

80 difference, in terms of mean and dispersion [*Cosh et al.*, 2006, 2008; *Kornelsen and Coulibaly*, 2013]. It

81 is based on the assumption that spatial SM fields are stable in time, which is not always true [*Yee et al.*,

82 2016]. Triple collocation (TC) can also be used to estimate the representativeness of ground stations

83 [*Miralles et al.*, 2010; *Gruber et al.*, 2013; *Chen et al.*, 2016]. It requires 3 datasets and is very sensitive

84 to the independence between the errors and between the signals and the errors [*Yilmaz and Crow*, 2014].

85 Finally, the "inverse footprint" method [*Orlowsky and Seneviratne*, 2014; *Nicolai-Shaw et al.*, 2015]

86 simply evaluates the synchronism between surrounding stations.

87 The spatial representativeness of SM datasets may be different depending on the time scale. Studies at

88 country and continental extents showed that large and small time scales have large and small

89   representativeness areas, respectively [*Cayan and Georgakakos*, 1995; *Vinnikov et al.*, 1996; *Entin et*

90   *al.*, 2000]. *Entin et al.* [2000] identified two spatio-temporal scales: the small scale was of the order of

91   some tens of meters and few days and was due to local processes such as infiltration, precipitation and

92   drainage; the large scale was of the order of some hundreds of km and 2-3 months and was due to

93   climatic atmospheric forcing. The works of *Chaney et al.*[2014] and *Su & Ryu* [2015] have provided

94   similar conclusions for footprint extents. *Chaney et al.* [2014] showed that, in the Little River

95   catchment, large spatial scale factors (land cover and evapotranspiration) influence SM seasonal cycles,

96   while the small ones (soil texture) do not. Similarly, *Su & Ryu* [2015] have showed that the correlation

97   between point and large-support datasets (*in situ* and satellite) increases with the time scale. However, at

98   the view of the literature on triple collocation (TC), we ascertain an alternative interpretation about SM

99   seasonal scales. TC studies have usually considered that there exist significant differences between the

100 seasonal components or "climatologies" of ground and satellite/model datasets due to their different

101 spatial support sizes [*Gruber et al.*, 2016]. For this reason, TC studies have systematically detrended the

102 SM series for the seasonal component. To our knowledge, this apparent divergence between

103 interpretations of the seasonal SM component has not been addressed yet in the literature.

104 The evaluation of SM representativeness on a per-time scale basis requires separating the SM series in

105 time scales. Moving averages have been applied to separate the seasonality and trend components (large

106 time scales) from the anomaly series component (shorter time scales) [*Gruber et al.*, 2013; *Nicolai-*

107 *Shaw et al.*, 2015]. Although events are localized with precision in the anomaly series, these are still

108 affected by part of the seasonal component. Fourier analysis has been used to analyze the power of each

109 time scale [*Katul et al.*, 2007; *Su et al.*, 2016], but it does not allow localizing events in time. More

110 advanced spectral techniques like the short-time Fourier transform and wavelet transforms can solve this

111 issue. Wavelet transforms have the advantage of localizing events in time with a precision that does not

112   depend on the time scale [*Barford et al.*, 1992; *Cornish et al.*, 2006]. Some examples of wavelets

113   applied to SM series include the study of daily to annual components at different depths [*Lauzon et al.*,

114   2004], the connections with other geophysical variables per time scale [*Graf et al.*, 2014] and the

115   correction of multiplicative and additive biases per time scale [*Su and Ryu*, 2015].

116   The objective of this study is to investigate the connection between spatial and time scales within

117   satellite footprints. The investigation is performed in three steps: first, a preliminary assessment of the

118   scales and their geophysical drivers is conducted on modeled SM data. Secondly, we investigate which

119   method is suited the best for assessing spatial representativeness (spatial scale) when SM time series are

120   decomposed in time scales. Time scales are obtained with wavelet transforms. The approaches tested for

121   assessing the spatial representativeness are temporal stability, triple collocation and two new ones: the

122   temporally correlated areas (CArea) method and an approach based on wavelet correlations (WCor). The

123   third and final step consists in analyzing actual measured SM data to verify the conclusions reached at

124   that point. To our knowledge, this is the first study of this kind to investigate the footprint extent with a

125   comprehensive set of methods and datasets. In addition, we analyze the seasonal components of point

126   and footprint-support series in order to solve the apparent divergence in literature mentioned before.

127   This article is structured as follows. Section 2 presents the methods used for the analyses in the time

128   domain (wavelets, section 2.1) and in the spatial domain (representativeness methods, section 2.2).

129   Section 3 describes the datasets. Section 4 gathers the results from each of the three steps of the

130   investigation in respectively three subsections. The conclusions are summarized in section 5.

131   **2   Materials and Methods**

132   **2.1   Time-Scale Decomposition of SM**

133   Wavelets are mathematical functions that can be used to decompose time series in a set of time scales

134   [*Foufoula-Georgiou and Kumar*, 1994; *Percival and Walden*, 2000]. Wavelet transforms are time-

135   frequency transforms: they detect the frequency components of the signal and also when events occur in

136   time. The continuous wavelet transform (CWT) is expressed as a collection of variables { $W(\tau,t) : \tau > 0$,

137    $-\infty < t < \infty$ }, where τ denotes the time scale (Eq. 1). It consists in convoluting the original signal $x(t)$

138   with a set of translated and stretched/shrinked versions of the wavelet basis function $\psi(t)$.

139
$$W(\tau, t) = \int_{-\infty}^{\infty} x(u)\psi\left(\frac{u-t}{\tau}\right) du \qquad \text{Eq. 1}$$

140   The maximal overlap discrete wavelet transform (MODWT) is a sub-sampled version of the CWT at

141   dyadic scales (Eq. 2).

142
$$\tau_j = 2^{j-1} T_s, \qquad j = 1, 2 \dots J0 \qquad \text{Eq. 2}$$

143   where J0 is the last level of decomposition, T_s is the sampling period of the original signal (in time

144   units), τ_j the time scale (in time units) and j the unit-less scale. The MODWT can be applied to any

145   sample size and is shift-invariant [*Percival & Walden*, 2000, pp. 159, 160].

146   The wavelet transform produces J0 series of *wavelet* coefficients $\{W_j(t)\}$ for the scales $\{\tau_j\}$ (j =

147   1, 2 ... J0) and one series of *scale* coefficients $V_{J0}(t)$ that contains all variations at scales larger than

148   τ_{J0}. For the sake of clarity, the scale series are usually referred as V_{J0} instead of $W_{(J0-\infty]}$. The inverse

149   transform of the Wj and V$_{J0}$ coefficients produces the *detail* (D$_j$) and *smooth* (S$_{J0}$) series, respectively.

150   The detail series represent anomalies at scale τ_j (rapid variations), i.e. differences in weighted averages

151   of periods of length τ_j or *slightly longer* [*Percival & Walden,* 2000, pp. 11, 59]. They are zero mean by

152   construction. The smooth series contain the remaining variations and the bias for time scales larger than

153   J0 (slow variations). The sum of the detail and smooth series recovers the original time series (Eq. 3).

154
$$x(t) = \sum_{j=1}^{J0} D_j + S_{J0} \qquad \text{Eq. 3}$$

155   One of the critical aspects of wavelet analysis is the choice of the maximum level of decomposition J0

156   and the wavelet basis function $\psi(t)$. The largest time scale at J0 should be smaller than the length of the

157   series $(2^{J0-1} < N)$. In this study, we use 6-month and 2-year series with a sampling period (T_s) of half

158   a day. Therefore, we select J0 = 8 for the 6-month series and J0 = 9 for the 2-year series. The list of

159   possible scales is provided in Table 1. Regarding the wavelet basis function, we will use the

160   Daubechies-4 (D4) [*Daubechies*, 1992] and the Haar [*Haar*, 1910] wavelet. While D4 better isolates

161   time scales due to its sharper response in the frequency domain, it is longer in time than Haar. Given that

162   the length of the wavelet at scale J0 should be shorter than the length of the series [*Cornish et al.*, 2006],

163   we select Haar for the 6-month series, and D4 for the 2-year series.

164   **2.2   Spatial Representativeness Metrics**

165   In this section, we describe the methods we use to evaluate the spatial representativeness: two existing

166   methods, temporal stability (TStab) and triple collocation (TC), and two new methods, the temporally-

167   correlated area (CArea) and the wavelet-based correlation (WCor). CArea is designed to serve as the

168   reference when working with modeled spatial fields since it accounts for all the local supports contained

169   within the footprint. In the case of dense *in situ* networks, the spatial sampling is insufficient to ensure

170   accurate CArea results. WCor is designed to serve as an alternative method to TStab and TC that, as will

171   be shown, require quite restrictive conditions constraining their use to limited range of time scales

172   **2.2.1   Temporal Stability (TStab)**

173   TStab was introduced by *Vachaud et al.* [1985] and has been thoroughly detailed in a number of

174   publications [*Martínez-Fernández and Ceballos*, 2005; *Cosh et al.*, 2006; *Mittelbach and Seneviratne*,

175 2012]. In short, TStab evaluates how the relative differences (RD$_i$, Eq. 4) between the spatial average

176 values SM$_{avg}$ and point-support values SM$_{pt-i}$ at the location i vary in time. The most representative point

177 time series is the one with both smaller mean RD (MRD$_i$, Eq. 5) and smaller standard deviation of RD

178 (SDRD$_i$, Eq. 6). In this study, stations with small and big MRD also had small and big SDRD,

179 respectively (not shown here). Thus, for the sake of concision, we bring the two metrics into one, the

180 RMSE$_i$ (Eq. 7), following the notation of *Jacobs et al.*, [2004].

181
$$RD_i(t) = \frac{SM_i(t) - SM_{avg}(t)}{SM_{avg}(t)}$$
Eq. 4

182
$$MRD_i = \frac{1}{N}\sum_{t=1}^{N} RD_i(t)$$
Eq. 5

183
$$SDRD_i = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(RD_i(t) - MRD_i)^2}$$
Eq. 6

184
$$RMSE_i = \sqrt{MRD_i^2 + SDRD_i^2}$$
Eq. 7

185 **2.2.2 Triple Collocation (TC)**

186 Triple collocation (TC) is a technique for estimating the random errors of three collocated datasets that

187 are meant to represent the same geophysical variable [*Stoffelen*, 1998]. It relies on a linear error model

188
$$x_k(t) = \alpha_k + \beta_k \theta(t) + \varepsilon_k(t)$$
Eq. 8

189 where k denotes one of the three datasets, α_k and β_k are calibration constants, $\theta(t)$ is the (unknown)

190 true SM and ε_k is the error term. In the case of SM, when TC is used to evaluate the spatial

191 representativeness, the TC triplet is formed by the *in situ* dataset (which is assessed for

192 representativeness) and two datasets of equivalent large supports. Supposing that the latter show

193 stronger similarities because of their similar support sizes and that the systems errors are much smaller

194 than the differences due to the spatial scale mismatch, the error metrics of the *in situ* dataset should

195 mainly reflect its spatial representativeness [*Vogelzang and Stoffelen*, 2012; *Gruber et al.*, 2016].

196 Two TC metrics are typically used, the variance of the unknown errors $\sigma^2_{\varepsilon_k}$ [*Miralles et al.*, 2010;

197 *Gruber et al.*, 2013] and the correlation between the dataset $\rho_{x_k,true}$ and the true soil moisture [*McColl*

198 *et al.*, 2014; *Chen et al.*, 2016]. In this study we use the TC-correlation coefficient because, unlike the

199 error variance, it is normalized by the total signal power and so allows the direct comparison of results

200 from different stations and networks.

201 Assuming that the covariances between the signal $\theta(t)$ and the errors $\varepsilon_k(t)$ and between the errors of

202 different datasets are null, the error variance and the TC-correlation estimators can be derived [*Chen et*

203 *al.*, 2016] and written as

204 $$\sigma^2_{\varepsilon_k} = \sigma^2_k - \sigma_{kl}\sigma_{km}/\sigma_{ml} \qquad \text{Eq. 9}$$

205 $$\rho_{x_k,true} = \pm\sqrt{\frac{\sigma_{kl}\sigma_{km}}{\sigma^2_k\sigma_{ml}}} \qquad \text{Eq. 10}$$

206 where σ^2_k is the variance of dataset k and σ_{kl}, σ_{km}, σ_{ml} are the cross-covariances between the two

207 datasets specified in the subscript. The 3 following conditions are necessary to compute Eq. 10 [*Chen et*

208 *al.*, 2016]: a) non-negative cross-correlation between all datasets; b) non-negative $\sigma^2_{\varepsilon_k}$; c) non-negative

209 $\rho^2_{x_k,true}$.

210 **2.2.3   Temporally Correlated Areas (CArea)**

211 *Nicolai-Shaw et al.* [2015] and *Orlowsky & Seneviratne* [2014] introduced the notion of "inverse

212 footprint" for *in situ* SM series that they define as the area surrounding a station where other stations

213 exhibit temporal similarity (correlation) above a specified threshold. In this study, we propose a

214 modification that we call the *temporally correlated areas* (CArea) method. The 3 main changes are:

215   a) It is only applied to SM gridded data. Even in the case of dense *in situ* networks the spatial

216   sampling is too sparse for detailed spatio-temporal analyses.

217   b) Pearson correlation replaces Spearman correlation, in order to be consistent with the other

218   approaches used that rely on the Pearson statistic.

219   c) The final metric is the percentage of pixels above a specific correlation threshold. The

220   mathematical formula is presented in Eq. 11, where $i0$ is the location where representativeness is

221   evaluated, M the number of locations i within the area A, $R_{x_i, x_{i_0}}$ the correlation between the time

222   series at locations i and $i0$, R_{th} the correlation threshold, and H the Heaviside function that is 0

223   and 1 for negative and positive numbers, respectively.

224 $$CArea(i_0) = \frac{1}{M} \sum_{\forall i \neq i_0 \subset A} H\left(R_{x_i, x_{i_0}} - R_{th} \right) \times 100 \ (\%) \qquad \text{Eq. 11}$$

225 The CArea method consists in calculating the percentage of time series within the study area that

226 correlate with the reference series x_{i0} above a specific threshold. The higher the percentage (and the

227 correlation threshold), the more representative is a location $i0$.

228 **2.2.4 Wavelet-based Correlation (WCor)**

229 The wavelet-based correlation (WCor) evaluates the representativeness of a location $i0$ on a per time-

230 scale basis. First, the point time series and the large-support series at that location are decomposed in

231 detail series with wavelet transforms. Then, correlation R_j between the detail series at each scale j is

232 computed:

233 $$R_j = R\{D_j^{point-i0}, D_j^{large}\}, \quad j = 1, 2 \dots J0 \quad \text{Eq. 12}$$

234 The WCor values are simply a measure of linear matching. They cannot by themselves quantify

235 separately the errors in the datasets and the spatial scale mismatch. However, the analysis of a collection

236 of *in situ* and modeled SM series in the following sections will show that they serve to understand the

237 connection between spatial and temporal scales and to compare the relative representativeness levels of

238 a set of stations.

3 Datasets

3.1 Local-support Datasets

241 We have selected for analysis the Little Washita watershed in USA [*Cosh et al.*, 2006] and the Yanco

242 area of the Murrumbidgee Soil Moisture Monitoring Network in Australia [*Smith et al.*, 2012]. They are

243 selected because both are monitored by dense *in situ* networks and have contrasting climatic conditions

244 (sub-humid and semi-arid, respectively) and SM forcing (irrigation is present in Yanco but not in Little

245 Washita). Little Washita will be used for the analysis of point *in situ* and modeled series and Yanco for

246 the analysis of time series of modeled SM gridded data. As explained in the introduction, modeled data

247 will serve for illustrating the connections between spatial scales, time scales and geophysical variables,

248 and actual measured data will be used for verifying the findings.

249 The Little Washita network has an extent of ~610 km^2. The average annual rainfall is 750 mm and most

250 of it takes place in spring and autumn [*Allen and Naney*, 1991]. The area is mainly covered by rangeland

251 and crops, soil texture is diverse (sands, loams and clays) and the topography is moderately rolling. The

252 network is made up of 20 permanent Stevens Hydra Probe stations installed at a depth of 5 cm with a

253 sensing range between 3 and 7 cm.

254 The Yanco network has an extent of ~3000 km^2. The average annual rainfall is around 400 mm with

255 precipitations concentrated in winter and spring. The area is mainly flat and is covered by pastures and

256 both dry and irrigated crops. The network is made up of 13 permanent Stevens Hydra Probe stations

257 providing SM integrated over the top 5 cm of soil.

258 **3.1.1 Modeled Series (Little Washita)**

259 The model in *Pan et al.* [2003] was specifically designed for the Little Washita network. Simplicity is its

260 major asset and the output SM time series are adequately realistic for our purposes. Three components

261 with distinctive temporal scales control the SM series dynamics: precipitation (short-term), texture

262 (short-to-medium term) and vegetation (seasonal term). These time scales could be represented by other

263 factors in other areas of study or in more complex models. For this reason, we consider that precipitation

264 includes irrigation, texture represents any soil memory process and vegetation represents any seasonal

265 signal, like temperature trends.

266 The model is summarized by the following equation:

267 $$SM(t) = \min\left\{\max\left\{SM(t-1) \cdot e^{\frac{-\eta(t) \cdot T_s}{Z}} + \frac{\gamma \cdot p(t)}{\eta(t) \cdot T_s} \cdot \left[1 - e^{\frac{-\eta(t) \cdot T_s}{Z}}\right], SM_{min}\right\}, SM_{max}\right\} \qquad \text{Eq. 13}$$

268 where T_s is the sampling period in hours (h), $p(t)$ is the cumulative precipitation (m) between $t-1$ and

269 t , γ is the interception by vegetation, η the loss coefficient (m/h) and Z is the penetration depth (m). The

270 loss coefficient is calculated as a function of the drainage coefficient *Ks* and the leaf-area index (*LAI*).

271 The parameters used in this study are described in

272 Table 2.

273 Two-year SM series are produced with this model at a 12 h sampling interval. Two sets of series (a, b)

274 are generated by varying the LAI amplitude. Each set is formed by a reference series (ref-a / ref-b) and

275 four sample series (1-, 2-, 3-, 4-a/b). The reference series are produced for a loam texture using the *in*

276 *situ* measured precipitation and the MODIS LAI time series observed at station #1. The ref-a time series

277 is shown in Figure 1-a, together with the true *in situ* series at station #1. Sample series are generated

278 identically to their respective reference series except for one variable (

279 Table 3): For sample series 1-a/b, we introduced random variations in precipitation amplitudes. For

280 sample series 2-a/b, we changed also the synchronization of some precipitation events (10 % of the

281 events are randomly selected and shifted in time by +0.5 day and another 10 % by -0.5 day). For series

282 3-a/b, we changed the texture to sand. Finally, for series 4-a/b, we introduced a 30 day time shift in the

283 seasonal component. The detailed setup is provided in

284   Table 3, where the variable changes are highlighted in italics.

3.1.2   *In Situ* Series (Little Washita)

286   The 20 *in situ* series of Little Washita for the 2012/07 - 2014/07 2-year period are selected. The data was

287   provided by the team of the U.S. Department of Agriculture (USDA) in charge of maintaining the

288   network. Data access and contact details can be found in the USDA Agricultural Research Service

289   website (http://ars.mesonet.org/). The spatial average of all the station series and the time series

290   measured at station #1 are shown for illustration in Figure 1-a. Since wavelet transforms need regularly

291   sampled time series, big gaps (> 1 month) are filled by linear regression with the most similar station

292   series. The percentage of filled gaps with this method is ~5.7 % of the entire series. The remaining gaps,

293   which represent ~1.1 % of the samples, are filled with a discrete cosine transform (DCT) approach

294   [*Wang et al.*, 2012]. The advantage of DCT is that it uses the full series −and not just local information-

295   to estimate the missing data based on the signal spectrum.

3.1.3   *In-situ*-DISPATCH Gridded Data (Yanco)

297   In this study, SM maps at 1 km resolution are generated by disaggregating the spatial average of the SM

298   *in situ* Yanco time series. Yanco *in situ* data is available from the OzNet hydrological monitoring

299   network website (http://www.oznet.org.au/). The disaggregation method used is derived from the

300   operational version of the Disaggregation based on Physical And Theoretical scale Change

301   (DISPATCH) algorithm [*Merlin et al.*, 2012, 2013; *Molero et al.*, 2016]. Former validation studies of

302   DISPATCH over the Yanco region gave satisfactory results [*Merlin et al.*, 2012; *Malbéteau et al.*,

303   2016]. The algorithm was originally designed to improve the resolution of satellite SM datasets by using

304   temperature and vegetation data from optical/thermal sensors like MODIS. Note that in this study, we

305   replace the satellite SM by the Yanco *in situ* average series, so that the SM maps are as close as possible

306   to the ground reality. DISPATCH preserves the spatial average by construction. The dataset will be

307   called *in situ*-DISPATCH (*in situ*-DIS).

308   The Yanco *in situ*-DIS series are sampled at SMOS overpass times (approximately 6 a.m. and 6 p.m.).

309   Long periods of clouds reduced dramatically the availability of DISPATCH data during the Austral

310   winter; as an example, most of the *in situ*-DIS series at the stations locations presented long periods of

311   unavailability (1-2 months) and data gaps represented 50 % of the series in average. As a consequence,

312   we only consider the 6 months from 2014/09 to 2015/03, which contain both shorter periods (below 9

313   days) and lower percentages of unavailability (~23 %). Data gaps are filled with the DCT approach

314   [*Wang et al.*, 2012].

315   **3.2   Large-support Datasets**

316   **3.2.1   SMOS**

317   The SMOS mission [*Kerr et al.*, 2001] is led by the European Space Agency (ESA) with collaboration

318   of the Centre National d'Etudes Spatiales (CNES) and the Centro Para el Desarrollo Tecnológico

319   Industrial (CDTI). The SMOS sensor is a passive 2D microwave interferometer observing the Earth at

320   L-band (1.4 GHz) dedicated for the observation of SM and ocean salinity. The mission provides SM

321   estimates in m^3/m^3 over the top ~5 cm surface layer. The footprint (support) has a resolution that varies

322   from 27 to 55 km depending on the observation geometry, with an average resolution of 43 km. The

323   maximum revisit time of SMOS is 3 days with crossing nodes at 6 a.m. and 6 p.m. local solar time for

324   ascending and descending orbits, respectively.

325   The SMOS data used in this study is obtained from the ESA Level-2 (L2) SM products (version 620).

326   The SM retrieval algorithm takes into account the landscape heterogeneity of the observed surface.

327   When the dominant land-cover is low-vegetated soil (like in this study), the brightness temperatures of

328   the low-vegetated part are modeled with the L-band Microwave Emission of the Biosphere (L-MEB)

329   forward model [*Wigneron et al.*, 2007]. Details of the L2 SM retrieval algorithms can be found in *Kerr*

330   *et al.* [2012, 2014].

331   The L2 grid nodes that are in the center of each *in situ* network are selected: the node 226157 for Little

332   Washita and the node 8174767 for Yanco. These are depicted in Figure 2, together with the position of

333   the ground stations of each network. Ascending and descending orbits are merged in one single time

334   series with a 0.5-day sampling period. The original SMOS time series for the Little Washita network is

335   shown in Figure 1-b. SM retrievals with probability of radio-frequency interference (RFI) higher than

336   10 % and data quality index (DQX) higher than 20 % are removed. The gaps represent 59 % of the Little

337   Washita and Yanco SMOS series and are evenly distributed: the mean number of consecutive gaps is 2.8

338   and the mean number of consecutive samples (without gaps) is 2.2. They are filled with the DCT

339   method.

340   **3.2.2   AMSR2**

341   The Advanced Microwave Scanning Radiometer 2 (AMSR2) is a passive multi-band scanning

342   radiometer onboard the Global Change Observation Mission Water 1 (GCOM-W1) satellite, launched

343   by the Japan Aerospace Exploration Agency (JAXA) in May 2012. Its revisit time is 1-2 days with

344   crossing nodes at 1:30 p.m. and 1:30 a.m. local solar time for ascending and descending orbits,

345   respectively. Since SM derived from lower frequencies is expected to be more accurate, the lowest

346   AMSR2 band (6.9 GHz, C-band) is selected here. At this frequency, the footprint is ~35 x 61 km (along

347   scan x along track) [*JAXA*, 2013] and the derived SM products represent the soil moisture of the top ~1–

348   2 cm surface layer.

349 Several AMSR2 SM products exist. We use the Land Parameter Retrieval Model (LPRM) products

350 [*Owe et al.*, 2008]. LPRM considers the surface as homogeneous within the footprint in terms of

351 vegetation scattering albedo, surface roughness, etc.. The product distributed by the NASA Goddard

352 Earth Sciences Data and Information Services Center showed unusual temporal patterns and positive

353 biases [*Cho et al.*, 2017], so we use an AMSR2-LPRM SM dataset directly provided by Dr. Parinussa.

354 We only LPRM products from descending overpasses (1:30 a.m.). They have been proved as more

355 accurate [*Draper et al.*, 2009; *Lei et al.*, 2015] than their ascending counterparts, probably due to the

356 more uniform surface temperature and soil moisture vertical profiles. For each network in this study,

357 the AMSR2 pixel closer to the selected SMOS node is chosen (Figure 2). The AMSR2 time series for

358 the Little Washita network is shown in Figure 1-b. SM estimates are discarded when either they are

359 equal to zero or when the quality mask values are higher than 68. On average, gaps represent 70 % of

360 the AMSR2 series and are uniformly distributed along the Little Washita and Yanco series: the mean

361 number of consecutive gaps is 3.8 and 1.9, respectively, and the mean number of consecutive samples is

362 1. Data gaps are filled with the DCT method.

363 **3.2.3   ECMWF**

364 We use the ECMWF SM dataset used by the SMOS L2 processor as initial guess in the retrieval loop.

365 This custom ECMWF dataset is obtained from the top 0-7 cm soil layer of the ECMWF forecast system

366 and has been interpolated in space and time to match the SMOS L2 grid and overpass times. The custom

367 ECWMF product is extracted from the SM_Init_Val field of the Level 2 Soil Moisture Data Analysis

368 Product (MIR_SMDAP2), which is available through the ESA SMOS dissemination web service

369 (https://smos-ds-02.eo.esa.int/oads/access/). More information on the ECMWF auxiliary product can be

370 found in *Kerr et al.* [2012, 2014, 2016]. The ECMWF time series for the Little Washita network in

371 shown in Figure 1-b. On average, gaps represent 48 % of the series and are uniformly distributed: the

372 mean number of consecutive gaps is 2.5 and the mean number of consecutive samples (without gaps) is

373 2.7. Gaps are filled with the DCT method.

4 Results and Discussion

4.1 Connection between Spatial and Time Scales

376 This section presents the first step of our investigation and seeks to reveal the existing connections

377 between spatial scales, time scales and geophysical drivers in SM modeled datasets. We analyze how

378 SM time scales are influenced by differences in the sources of SM variability (forcing events, soil

379 memory and seasonal sources), for which the Little Washita modeled series were specifically designed.

380 To this end, we evaluate the correlation between each sample series and its respective reference series (*a*

381 or *b*, Table 3) on a per-time scale basis.

382 The correlation of each sample-reference series pair is depicted in Figure 3 with differently colored

383 lines. Solid lines correspond to pairs of the *a* group and dashed lines to the *b* group. Differences in

384 forcing events (blue and red lines) deteriorate the correlation, at least in the first time scales (≤ 2 days).

385 Moreover, de-synchronizations produce irregular correlation patterns (red lines). Regarding texture

386 heterogeneity (magenta lines), it deteriorates the correlations of middle scales up to the first seasonal

387 scale (32 to 64-day scales). This illustrates that both meteorological forcing and surface memory can

388 contribute to the month and seasonal scale signatures. Finally, when the seasonal component is not

389 synchronized, the correlation at month and seasonal scales is hampered. This happens only when the

390 seasonal component represents an important part of signal (case 4-a), otherwise, the correlation is

391 maintained (case 4-b).We have just shown the connections between some of the sources of SM

392 variability and SM time scales, from a model perspective. Do these sources also exhibit characteristic

393   spatial scales? For an exhaustive spatial investigation, the area under study requires to be fully sampled,

394   so in the next experiment we use the time series of *in situ*-DIS spatialized data. We evaluate the spatial

395   representativeness of the pixels containing an *in situ* station on a per time-scale basis, by applying

396   CArea to their wavelet decomposed series. We also evaluate the representativeness of two other series

397   that are expected to represent the satellite footprint better than the point *in situ* series: the field average

398   series (FAvg, the average of all the pixels) and the network average series (NAvg, the average of the

399   pixels containing a station).

400   The results are presented in Figure 4, where each line represents the CArea values obtained for a specific

401   pixel for a range of correlation thresholds. Regardless the FAvg and the NAvg series that have their own

402   names, the ID of each pixel corresponds to the number of the *in situ* station contained within. We

403   observe that the lines move to the right and are more distant from each other as the time scale increases.

404   This implies that, in general, spatial representativeness increases with the time scale, but the evolution is

405   not the same for all locations. The latter could be explained by the combination between the propagation

406   of small scale effects and the appearance of larger scale SM factors (the propagation of small scale

407   effects was shown in the modeled Little Washita series). We also notice that the field and the network

408   average series are the most representative ones at all time scales.

409   Small time scales (0.5-2 days) exhibit the smallest correlated area, with less than 25 % of the area

410   correlated above 0.5 (Figure 4). This can be due to three possible reasons: i) gap-filling, ii) noise from

411   the disaggregation method and inputs, iii) important spatial heterogeneity. In order to assess the impact

412   of gap-filling, we used measured *in situ* series, where we could set the same gaps as those in the *in situ*-

413   DIS series and compare scores before and after filling the gaps. Since the number of spatial samples was

414   not large enough for applying CArea (13 stations), we simply computed the wavelet correlation scores.

415   We found that, at the 0.5-2 day scales, correlation decreased by 0.08 on average. This means that gap-

416   filling does not change significantly the first 3 time scales as shown in Figure 4, with respect to the other

417   time scales. Concerning the disaggregation noise we expect it to be negligible with respect to the

418   geophysical heterogeneity because the analysis of actual *in situ* series, as it will be shown in section 4.3,

419   exhibited similar decorrelation levels. Hence, we may conclude that the spatial heterogeneity is the main

420   cause of low correlation at sub-weekly time scales. In this case, the heterogeneity is most likely

421   controlled by irrigation: the Yanco area contains both irrigated and non-irrigated plots and we showed

422   before that precipitation (and by extension, irrigation) produced de-correlation at short-time scales.

423   Regarding weekly scales (8-16 days), most of the series have more than 50 % of the surface correlated

424   above 0.5 and 0.6 correlation points, respectively. This suggests that there is little soil heterogeneity,

425   according to our analysis of Figure 3 where texture was associated to middle scales. The month scale (32

426   days) breaks the tendency of increasing representativeness with time scale. As we deducted from Figure

427   3, such drops in correlation appear at similar time scales when precipitation events are not synchronized,

428   a phenomenon that should be present in Yanco because of irrigation. Taking also that into account the

429   32-days scale has small temporal variance, similar to that of the 2-days scale for this dataset (not shown

430   here), the signal-to-noise ratio might be quite low and induce low correlation (as demonstrated by

431   [*Berger and Sweney*, 1965; *Goodwin and Leech*, 2006]).

432   The largest scales (16-64 days) deserve special attention. Firstly, we recall that the relative positions

433   between the lines change in Figure 4. This justifies the separate evaluation of spatial representativeness

434   per time scale. For example, location #1 is a good option if we are interested in seasonal changes (64-

435   days scale) but it is not for week-scale applications (8- and 16-days scales). Secondly and most

436   importantly, the seasonal 64-days scale is the scale that exhibits the largest areas with correlation very

437   close to 1: the most representative series exhibits ~40 % of the area with a correlation above 0.9 (Figure

438   4). However, there are some locations that have extremely small representativeness areas (#13) while

439   others have extremely large ones (#9, #10).

440   In order to investigate deeper, Figure 5 presents the correlation maps derived for the FAvg series, prior

441   to the calculation of the CArea percentages for this series. It shows the same overall trend of increasing

442   representativeness with time scale, including the correlation drop at the 32-day scale explained before. It

443   also corroborates that at the 64-day scale, locations can be either highly representative of footprint SM

444   (correlation close to 1, in yellow), or not at all (correlation < 0.5, in dark blue). Additional experiments

445   (not included here), showed that concurrent heterogeneities in precipitation synchronization and texture

446   affected seasonal time scales, which can explain the observed dispersion in representativeness. From

447   this, we conclude that the seasonal component of SM is made up of standalone seasonal elements

448   (vegetation growing cycle, temperature trends, etc.) along with the integration over time of smaller time

449   scale components, like short-time precipitation events and surface memory.

450   The results presented in this section solve the apparent opposition between the detrending in TC studies

451   and the conclusions in *Su and Ryu* [2015] about seasonal scales that was mentioned in the introduction.

452   Both Figure 4 and Figure 5 reveal that, at seasonal time scales, both situations coexist: some locations

453   exhibit important differences with respect to the footprint time series, as suggested in TC studies, but

454   also a large number of locations exhibit good synchronization, as proposed by *Su & Ryu* [2015]. Finally,

455   we have also shown that time and spatial scales are connected in the model-based Little Washita and

456   Yanco datasets. We hope to find similar behavior in actual *in situ* series (section 4.3), given that both

457   model datasets are dependent on measured *in situ* data.

458   **4.2   Inter-comparison of Methods for Spatial Representativeness Assessment**

459 Herein, we describe the second step of our investigation, which is dedicated to finding the best methods

460 for assessing spatial representativeness of SM datasets, especially when SM time series are decomposed

461 in time scales. To this end, we compare the performances of TStab, TC, CArea and WCor methods

462 when applied to the Yanco *in situ*-DIS dataset for the 09/2014 - 03/2015 6-month period. Because of the

463 CArea method, the area of study includes all the stations plus a 0.05° extension to avoid borders effects

464 in peripheral stations. The TC triplets are made up of one local-support dataset (one pixel *in situ*-DIS

465 series) and two large-support datasets (the SMOS dataset and either the AMSR2 or the ECMWF

466 dataset).

4.2.1 Original Series

468 Figure 6-a shows the spatial representativeness values obtained with each method on each selected

469 location (pixel). The vertical axis is oriented from small to large representativeness, from bottom to top.

470 Results are grouped per method: at the left, the CArea percentages; in the middle, the TC correlation

471 $\rho_{pixel,true}$ values; and at the right, the TStab RMSE values (in reverse vertical-axis order). Some

472 locations (markers) are missing from the TC groups because the preliminary test on the error variances

473 (section 2.2.2) gave a negative value. This can be due to temporal biases, which can cause an imbalance

474 between the dataset variance and the product of covariances (Eq. 9). TStab exhibits the largest

475 disagreements with respect to the other methods. In agricultural sites, human decisions (cropping,

476 irrigation) undermine TStab performances because they affect the temporal stability of the spatial

477 distribution of SM [*Yee et al.*, 2016].

478 In Figure 6-a, the ranking of the locations in terms of representativeness is not the same for CArea and

479 TC methods. Moreover, the values of the two TC variants are not coincident in general, although they

480 both assign the largest values to the network average and pixels 9 and 12. All these differences could be

481   induced by seasonal biases. Typically, TC studies removed the 30-days average component in order to

482   have more chances to fulfill TC requirements (e.g. [*Miralles et al.*, 2010; *Chen et al.*, 2016]). In our

483   case, we take advantage of the wavelet decomposition technique to provide a detrended triplet where

484   variations larger than 32 days are removed. Figure 6-b shows the representativeness scores of the

485   detrended series. The ordering of the locations is more similar for the two TC variants than in Figure 6-

486   a. The wavelet-based detrending is beneficial because AMSR2 was positively biased during the first half

487   of the period (not shown here). This can be due to C-band being more sensitive to vegetation and

488   atmospheric factors than L-band. However, detrending does not prevent the TC and the CArea methods

489   to provide very different results (Figure 6-b). They both agree in attributing more spatial

490   representativeness to the network average and locations #9, #5, #4, #10, #8, while smaller spatial

491   representativeness to locations #1, #2, #6, #7, but still some locations like #12, #13 and #3 exhibit large

492   discrepancies. This reveals that detrending improves TC performance but it does not succeed by itself to

493   ensure that TC conditions are perfectly fulfilled.

494   **4.2.2   Time-scale Decomposed Series**

495   The methods presented show significant differences in performance depending on whether some time

496   scales, especially the seasonal one, are removed or not. Herein, we study the phenomenon in more detail

497   at all time scales. The decomposition in time scales allows using the WCor approach, which compares

498   the series of the selected pixels with the series of their spatial average (NAvg), on a per-time scale basis

499   In Figure 7, each plot contains the representativeness scores obtained with the different methods,

500   including WCor, at a different time scale. There is a large absence of TC scores at the half-day, 1-day,

501   32- and 64-day scales. This is either because they are off vertical axis limits, or because they fail the TC

502   preliminary tests (e.g. most of the times the correlations between the datasets were too low, below 0.5,

503   not shown here).

504   In Figure 7, the relative ranking of TC values differ to a great extent from that of the WCor and CArea

505   values. Considering only the ECMWF-based configuration of the TC scores, the highest concordance

506   between TC and WCor rankings occurs at the 8-days scale. The mismatch at larger scales ($\geq$ 16 days)

507   can be explained because the number of independent samples is drastically reduced due to wavelet

508   filtering. The length of the series (366 samples) is already lower than TC recommendations: around 500

509   samples are needed for error variances $\sigma_{\varepsilon_k}^2$ estimated with low uncertainty (11 %) [*Zwieback et al.*,

510   2012]. The mismatch at small scales ($\leq$2 days) is probably due to the very low correlation between the

511   datasets, which hampers the validity of the linear model assumption. All this suggests that TC should be

512   applied neither to too short series nor to the shortest time scales.

513   Finally, the WCor and CArea methods give consistent results: the ranking of the locations is similar for

514   all time scales. This is significant since the fact that a location correlates well/badly with the rest of

515   locations (CArea) does not imply that it correlates well/badly with the network average (WCor), and

516   *vice-versa:* the correlation between the average and a point series cannot be simply summarized as the

517   average of point-to-point correlation values. From this we conclude that WCor is a robust method for the

518   evaluation of spatial representativeness on a per-time scale basis.

519   **4.3   Spatial Representativeness Assessment of *In Situ* Series**

520   This last section of results investigates whether the conclusions reached on modeled SM data apply to

521   true *in situ* series, concretely those of the Little Washita network. The CArea method will no longer be

522   applied since the spatial sampling is not sufficiently dense. The WCor method will be also tested on

523   other large-support datasets different from the *in situ* average (SMOS, AMSR2 and ECMWF). It will

524   allow exploring whether WCor could be applied to sparse networks (a single *in situ* station per

525   footprint). The 2012/07 - 2014/07 2-year period is selected.

526   The results of the WCor and TC methods are presented in Figure 8. In contrast to the *in situ*-DIS Yanco

527   case (section 4.2), much more TC scores are present, which is due to the larger number of samples.

528   Figure 8 confirms the connection between time and spatial scales described in section 4.1: spatial

529   representativeness increases with the time scale and the largest time scales (64 and 128 days) present the

530   largest scatter in representativeness values. However, a drop in representativeness scores appears at the

531   64-day scale and concerns all the method configurations except the WCor-*in situ* avg. The most likely

532   explanation is that the Little Washita network covers only half of the surface of the satellite footprints.

533   As a consequence, the *in situ* series should present similar differences with respect to the satellite

534   products in terms of precipitation and surface memory and, as explained in section 4.1, these elements

535   can cause decorrelation in the 32- and 64-day scales.

536   Is the gap-filling the root cause of the low representativeness scores at sub-weekly scales? According to

537   the previous section, the gap-filling in the point and average *in situ* series has a marginal effect. In

538   addition, in this section we evaluated its effect on the large scale datasets. The procedure consisted in

539   setting the large-scale datasets gaps in the *in situ* series and computing the scores again. In the case of

540   WCor, we observed that, at scales smaller than 4 days, the experiment induced a small reduction in

541   variance and an increase in correlation of between 0.05 and 0.2. According to these results, the gap-

542   filling does not change the relative scores presented in Figure 8 and in this study in general: the scores at

543   scales smaller than 4 days remain much lower than those of larger time scales, even after taking into

544   account the correlation increase due to gap-filling.

545   When TC and WCor approaches are compared, similarities are found by groups (Figure 8): ECMWF-

546   based TC results match well with the WCor results when the large-support dataset is either the *in situ*

547   average, SMOS or ECMWF (1st group), while AMSR2-TC values match well with the WCor-AMSR2

548   values (2nd group). This highlights that both TC and WCor methods have a high sensibility to the choice

549   of the large support dataset. Two main features can explain the differences between the first and the

550   second group. First, the gaps and the sensing time of AMSR2 series are different to those of the second

551   group. Secondly, the interpolation of the custom ECMWF dataset and its use as first guess in the SMOS

552   retrieval system could foster similarity with SMOS-based scores. However, at the view of recent

553   analyses of both products [*Kerr et al.*, 2016], the SMOS retrievals are independent of ECMWF values.

554   The TC-WCor consistency is lost at the 64-days scale for the first group and at the 32-days scale for the

555   second group. This is probably caused by a poorer performance of the TC method due to the reduction

556   in the number of independent samples along with a correlation decrease between the datasets at those

557   particular scales. This can be seen in Figure 9, where the correlation between the datasets is shown. We

558   also observe that the higher correlation values for the first group at the 128-days scale (Figure 9-a)

559   seems related to the good consistency between WCor and TC results at this scale in Figure 8. For

560   example, both methods designate stations #3, #11, #14, #15, #17, #19 as the most representative and

561   stations #2, #4, #5, #8, #16 and #20 as the least representative ones.

562   SMOS and ECMWF WCor results are the most similar to the *in situ*-avg WCor scores (Figure 8), so we

563   consider necessary to examine them in more detail. First, SMOS- and ECMWF-based correlations are

564   very low (< 0.5) at the first 3 scales (0.5-2 days) while the *in situ*-avg correlations are higher than 0.6. In

565   the case of satellite datasets, this could be due to high-frequency noise, but not in the case of models like

566   ECMWF that are governed by a smoother model structure. Another more likely explanation is related to

567   the spatial support: the spacing between the *in situ* stations is larger than the correlation length of SM,

568   which ranges between some meters to some hundreds of meters [*Western et al.*, 1998, 2004; *De Lannoy*

569   *et al.*, 2006]. As a consequence, the *in situ* average is computed with an undersampled surface, which

570   misses small spatial scale phenomena, while satellite sensors observe a continuous sampled surface.

571   Moreover, satellite sensors estimate SM from the energy integrated over the footprint, which is not

572   necessarily equal to the integral of SM due to non-linearities in the models and in the scaling of

573   parameters [*Crow et al.*, 2001; *Crosson et al.*, 2010]. From this we conclude that the smallest time scales

574   (≤ 2 days) are not good choices to validate satellite estimates given the large geophysical mismatch

575   between satellite and *in situ* measurements.

576   Regarding the middle scales (4-16 days), the ranking of ECMWF-WCor is more similar to the *in situ*-

577   avg ranking than the SMOS one (Figure 8), which we attribute to SMOS observational noise. However,

578   at last scales (32-128 days) we observe the opposite. Therefore, we consider SMOS as a good large-

579   support dataset to be used for spatial representativeness assessment in the Little Washita region,

580   especially at the month and seasonal scales.

581   **5   Conclusions**

582   Satellite surface SM products are often validated with ground samples by direct comparison, despite the

583   different spatial supports of the two datasets (~50 km and a few centimeters, respectively). Ground

584   samples can represent areas larger than their measurement support. The representativeness area may

585   vary with the time scale [*Entin et al.*, 2000]. This study sought to investigate the connections between

586   SM spatial and time scales within typical coarse scale satellite footprint-size areas. For this purpose, we

587   evaluated the spatial representativeness of different locations at a range of time scales with various

588   methods: triple collocation (TC), temporal stability analysis (TStab), the percentage of correlated area

589   (CArea) and a new proposed approach consisting in wavelet-based correlations (WCor).

590   The comparison of the four approaches revealed that TStab, although applicable to SM absolute values,

591   could not be applied to wavelet decomposed series because of their multiple zero-crossings. TC could

592   not give any results or gave results that were not consistent with the other methods under two situations:

593   at short time scales (0.5-2 days), because the correlation between the datasets was too low, and at larger

594   time scales (larger than 8 days in the case of 6-month series and larger than 32 days in the case of 2-year

595   series), because the number of independent samples was too low after wavelet filtering. CArea and

596   WCor results were consistent in general. WCor is less sensitive to the spatial sampling density than

597   CArea, so it is a robust method for *in situ* networks that moreover requires less restrictive conditions

598   than the 3 other approaches presented.

599   By applying TC, CArea and WCor to modeled and true *in situ* time series in the Little Washita

600   watershed and to spatialized SM data in the Yanco area, we found that SM spatial and time scales were

601   connected. The series were sampled every 0.5 days. Precipitation and irrigation were found responsible

602   of small representativeness areas at small time scales (0.5-2 days). As the time scale increased from 0.5

603   days to 128 days, the spatial representativeness scores tended to increase as well; however, they became

604   more scattered. This was explained by different geophysical factors. First, de-synchronizations in

605   precipitation were propagated to larger time scales preventing representativeness to regularly increase at

606   some locations. Secondly, we observed that the seasonal scale did not only include seasonal signals

607   (vegetation growth, temperature trends, etc.) but also the temporal integration of precipitation and soil

608   memory responses from short and medium time scales.

609   This is, to our knowledge, the first comprehensive investigation on the connection between SM spatial

610   and time scales within the satellite footprint (~50 km). It has revealed that time decompositions along

611   with the WCor method are promising tools for improving satellite validation and modeling of surface

612   soil moisture. At small time scales (below 4 days), the spatial scale mismatch between satellite/model

613   series (SMOS, AMSR2, ECMWF) and *in situ* series was found extremely large and similar for all

614   stations. Therefore, we suggest not taking into account these time scales in the validation of satellite

615   products. At the seasonal scale, some locations were observed very similar to the footprint-support

616   series, while some others were very different. This explained why in some previous studies seasonal

617 scales were found similar for both *in situ* and satellite series (e.g. [*Su and Ryu*, 2015]) while in TC

618 studies they were supposed intrinsically different so seasonal detrending was applied (e.g. [*Gruber et al.*,

619 2016]). Finally, the findings of this study can contribute to other SM applications like downscaling or

620 modeling: multi-scale algorithms can be built based on the specific interactions at each time and spatial

621 scale. Given its time-scale dependence, spatial variability should be addressed differently depending on

622 the time scale.

631 **Bibliography**

632 Allen, P. B., and J. W. Naney (1991), *Hydrology of the Little Washita River Watershed, Oklahoma:*

633   *Data and Analyses*, United States Department of Agriculture, Agricultural Research Service, ARS-

634   90.

635 Barford, L. A., R. S. Fazzio, and D. R. Smith (1992), *An introduction to wavelets*, Hewlett-Packard

636   Company.

637   Berger, R. A., and A. B. Sweney (1965), Variance and Correlation Coefficients, *Res. Quarterly. Am.*

638   *Assoc. Heal. Phys. Educ. Recreat.*, *36*(3), 368–370, doi:10.1080/10671188.1965.10614705.

639   Bircher, S., N. Skou, K. H. Jensen, J. P. Walker, and L. Rasmussen (2012), A soil moisture and

640   temperature network for SMOS validation in Western Denmark, *Hydrol. Earth Syst. Sci.*, *16*(5),

641   1445–1463, doi:10.5194/hess-16-1445-2012.

642   Al Bitar, A., D. J. Leroux, Y. H. Kerr, O. Merlin, P. Richaume, A. K. Sahoo, and E. F. Wood (2012),

643   Evaluation of SMOS Soil Moisture Products Over Continental U.S. Using the SCAN / SNOTEL

644   Network, *IEEE Trans. Geosci. Remote Sens.*, *50*(5), 1572–1586.

645   Blöschl, G., and M. Sivapalan (1995), Scale issues in hydrological modelling: A review, *Hydrol.*

646   *Process.*, *9*(September 1994), 251–290, doi:10.1002/hyp.3360090305.

647   Cayan, D. R., and K. P. Georgakakos (1995), Hydroclimatology of Continental Watersheds: 2. Spatial

648   Analyses, *Water Resour. Res.*, *31*(3), 677–697, doi:10.1029/94WR02376.

649   Chaney, N. W., J. K. Roundy, J. E. Herrera-Estrada, and E. F. Wood (2014), High-resolution modeling

650   of the spatial heterogeneity of soil moisture: Applications in network design, *Water Resour. Res.*,

651   1–20, doi:10.1002/2013WR014964.

652   Chen, F., W. T. Crow, P. J. Starks, and D. N. Moriasi (2011), Improving hydrologic predictions of a

653   catchment model via assimilation of surface soil moisture, *Adv. Water Resour.*, *34*(4), 526–536,

654   doi:10.1016/j.advwatres.2011.01.011.

655   Chen, F., W. T. Crow, A. Colliander, M. H. Cosh, T. J. Jackson, R. Bindlish, R. H. Reichle, S. K. Chan,

656   D. D. Bosch, P. J. Starks, D. C. Goodrich, and M. S. Seyfried (2016), Application of Triple

657   Collocation in Ground-Based Validation of Soil Moisture Active/Passive (SMAP) Level 2 Data

658   Products, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 1–14,

659   doi:10.1109/JSTARS.2016.2569998.

660   Cho, E., C.-H. Su, D. Ryu, H. Kim, and M. Choi (2017), Does AMSR2 produce better soil moisture

661   retrievals than AMSR-E over Australia?, *Remote Sens. Environ.*, *188*, 95–105,

662   doi:10.1016/j.rse.2016.10.050.

663   Colliander, A., T. J. Jackson, R. Bindlish, S. K. Chan, et al. (2017), Validation of SMAP surface soil

664   moisture products with core validation sites, *Remote Sens. Environ.*, *191*, 215–231,

665   doi:10.1016/j.rse.2017.01.021.

666   Cornish, C. R., C. S. Bretherton, and D. B. Percival (2006), Maximal overlap wavelet statistical analysis

667   with application to atmospheric turbulence, *Boundary-Layer Meteorol.*, *119*(2), 339–374,

668   doi:10.1007/s10546-005-9011-y.

669   Cosh, M. H., T. J. Jackson, P. J. Starks, and G. Heathman (2006), Temporal stability of surface soil

670   moisture in the Little Washita River watershed and its applications in satellite soil moisture product

671   validation, *J. Hydrol.*, *323*(1–4), 168–177, doi:10.1016/j.jhydrol.2005.08.020.

672   Cosh, M. H., T. J. Jackson, M. S. Moran, and R. Bindlish (2008), Temporal persistence and stability of

673   surface soil moisture in a semi-arid watershed, *Remote Sens. Environ.*, *112*(2), 304–313,

674   doi:10.1016/j.rse.2007.07.001.

675   Crosson, W. L., A. S. Limaye, and C. A. Laymon (2010), Impacts of spatial scaling errors on soil

676   moisture retrieval accuracy at L-band, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, *3*(1), 67–

677   80, doi:10.1109/JSTARS.2010.2041636.

678   Crow, W. T., M. Drusch, and E. F. Wood (2001), An observation system simulation experiment for the

679   impact of land surface heterogeneity on AMSR-E soil moisture retrieval, *IEEE Trans. Geosci.*

680   *Remote Sens.*, *39*(8), 1622–1631, doi:10.1109/36.942540.

681   Crow, W. T., A. A. Berg, M. H. Cosh, A. Loew, B. P. Mohanty, R. Panciera, P. de Rosnay, D. Ryu, and

682   J. P. Walker (2012), Upscaling sparse ground-based soil moisture observations for the validation of

683   coarse-resolution satellite soil moisture products, *Rev. Geophys.*, *50*(2), 1–20,

684   doi:10.1029/2011RG000372.

685   Daly, E., and A. Porporato (2005), A Review of Soil Moisture Dynamics: From Rainfall Infiltration to

686   Ecosystem Response, *Environ. Eng. Sci.*, *22*(1), 9–24, doi:10.1089/ees.2005.22.9.

687   Daubechies, I. (1992), 6. Orthonormal Bases of Compactly Supported Wavelets, in *Ten Lectures on*

688   *Wavelets*, pp. 167–213, Society for Industrial and Applied Mathematics.

689   Douville, H. (2004), Relevance of soil moisture for seasonal atmospheric predictions: Is it an initial

690   value problem?, *Clim. Dyn.*, *22*(4), 429–446, doi:10.1007/s00382-003-0386-5.

691   Draper, C. S., J. P. Walker, P. J. Steinle, R. A. M. de Jeu, and T. R. H. Holmes (2009), An evaluation of

692   AMSR-E derived soil moisture over Australia, *Remote Sens. Environ.*, *113*(4), 703–710,

693   doi:10.1016/j.rse.2008.11.011.

694   Draper, C. S., R. H. Reichle, G. J. M. De Lannoy, and Q. Liu (2012), Assimilation of passive and active

695   microwave soil moisture retrievals, *Geophys. Res. Lett.*, *39*(4), doi:10.1029/2011GL050655.

696   Drusch, M. (2007), Initializing numerical weather prediction models with satellite-derived surface soil

697   moisture: Data assimilation experiments with ECMWF's integrated forecast system and the TMI

698   soil moisture data set, *J. Geophys. Res. Atmos.*, *112*(3), doi:10.1029/2006JD007478.

699   Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai (2000), Temporal and

700   spatial scales of observed soil moisture variations in the extratropics, *J. Geophys. Res.*, *105*(D9),

701   865–877, doi:10.1029/2000JD900051.

702   Foufoula-Georgiou, E., and P. Kumar (1994), *Wavelets in Geophysics, Volume 4*, Academic Press, San

703   Diego, California, USA.

704   Goodwin, L. D., and N. L. Leech (2006), Understanding Correlation: Factors That Affect the Size of r,

705   *J. Exp. Educ.*, *74*(3), 249–266, doi:10.3200/JEXE.74.3.249-266.

706   Graf, A., H. R. Bogena, C. Drüe, H. Hardelauf, T. Pütz, G. Heinemann, and H. Vereecken (2014),

707   Spatiotemporal relations between water budget components and soil water content in a forested

708   tributary catchment, *Water Resour. Res.*, *50*(6), 4837–4857, doi:10.1002/2013WR014516.

709   Gruber, A., W. A. Dorigo, S. Zwieback, A. Xaver, and W. Wagner (2013), Characterizing Coarse-Scale

710   Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture

711   Network, *Vadose Zo. J.*, *12*(2), 1–16, doi:10.2136/vzj2012.0170.

712   Gruber, A., C.-H. Su, S. Zwieback, W. T. Crow, W. A. Dorigo, and W. Wagner (2016), Recent

713   advances in (soil moisture) triple collocation analysis, *Int. J. Appl. Earth Obs. Geoinf.*, *45*, 200–

714   211.

715   Haar, A. (1910), Zur Theorie der orthogonalen Funktionensysteme, *Math. Ann.*, *69*(3), 331–371,

716   doi:10.1007/BF01456326.

717   Jackson, T. J., M. H. Cosh, R. Bindlish, P. J. Starks, D. D. Bosch, M. S. Seyfried, D. C. Goodrich, M. S.

718   Moran, and J. Du (2010), Validation of advanced microwave scanning radiometer soil moisture

719   products, *IEEE Trans. Geosci. Remote Sens.*, *48*(12), 4256–4272,

720   doi:10.1109/TGRS.2010.2051035.

721   Jacobs, J. M., B. P. Mohanty, E.-C. Hsu, and D. Miller (2004), SMEX02: Field scale variability, time

722   stability and similarity of soil moisture, *Remote Sens. Environ.*, *92*(4), 436–446,

723   doi:10.1016/j.rse.2004.02.017.

724   JAXA (2013), *GCOM-W1 "SHIZUKU" Data Users Handbook*.

725   Katul, G. G., A. Porporato, E. Daly, A. C. Oishi, H. Kim, P. C. Stoy, J. Juang, and M. B. Siqueira

726   (2007), On the spectrum of soil moisture from hourly to interannual scales, , *43*, 1–10,

727   doi:10.1029/2006WR005356.

728   Kerr, Y. H., P. Waldteufel, J.-P. Wigneron, J. M. Martinuzzi, J. Font, and M. Berger (2001), Soil

729   moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE Trans.*

730   *Geosci. Remote Sens.*, *39*(8), 1729–1735, doi:10.1109/36.942551.

731   Kerr, Y. H., P. Waldteufel, P. Richaume, J.-P. Wigneron, P. Ferrazzoli, A. Mahmoodi, A. Al Bitar, F.

732   Cabot, C. Gruhier, S. E. Juglea, D. J. Leroux, A. Mialon, and S. Delwart (2012), The SMOS Soil

733   Moisture Retrieval Algorithm, *Geosci. Remote Sens.*, *50*(5), 1384–1403.

734   Kerr, Y. H., P. Waldteufel, P. Richaume, P. Ferrazzoli, and J.-P. Wigneron (2014), *SMOS Level-2*

735   *Processor Soil Moisture Algorithm Theoretical Basis Document (ATBD)*, SM-ESL (CBSA),

736   Toulouse.

737   Kerr, Y. H., A. Al-Yaari, N. J. Rodríguez-Fernández, M. Parrens, et al. (2016), Overview of SMOS

738   performance in terms of global soil moisture monitoring after six years in operation, *Remote Sens.*

739   *Environ.*, *180*, 40–63, doi:10.1016/j.rse.2016.02.042.

740   Kornelsen, K. C., and P. Coulibaly (2013), McMaster Mesonet soil moisture dataset: Description and

741   spatio-temporal variability analysis, *Hydrol. Earth Syst. Sci.*, *17*(4), 1589–1606, doi:10.5194/hess-

742   17-1589-2013.

743   De Lannoy, G. J. M., N. E. C. Verhoest, P. R. Houser, T. J. Gish, and M. Van Meirvenne (2006), Spatial

744   and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE3),

745   *J. Hydrol.*, *331*(3–4), 719–730, doi:10.1016/j.jhydrol.2006.06.016.

746   Lauzon, N., F. Anctil, and J. Petrinovic (2004), Characterization of soil moisture conditions at temporal

747   scales from a few days to annual, *Hydrol. Process.*, *18*(17), 3235–3254, doi:10.1002/hyp.5656.

748   Legates, D. R., R. Mahmood, D. F. Levia, T. L. DeLiberty, S. M. Quiring, C. Houser, and F. E. Nelson

749   (2011), Soil moisture: A central and unifying theme in physical geography, *Prog. Phys. Geogr.*,

750   *35*(1), 65–86, doi:10.1177/0309133310386514.

751   Lei, F., W. T. Crow, H. Shen, R. M. Parinussa, and T. R. H. Holmes (2015), The impact of local

752   acquisition time on the accuracy of microwave surface soil moisture retrievals over the contiguous

753   United States, *Remote Sens.*, *7*(10), 13448–13465, doi:10.3390/rs71013448.

754   Loew, A., and F. Schlenz (2011), A dynamic approach for evaluating coarse scale satellite soil moisture

755   products, *Hydrol. Earth Syst. Sci.*, *15*(1), 75–90, doi:10.5194/hess-15-75-2011.

756   Malbéteau, Y., O. Merlin, B. Molero, C. Rüdiger, and S. Bacon (2016), DisPATCh as a tool to evaluate

757   coarse-scale remotely sensed soil moisture using localized in situ measurements: Application to

758   SMOS and AMSR-E data in Southeastern Australia, *Int. J. Appl. Earth Obs. Geoinf.*, *45*, *Part B*,

759   221–234, doi:10.1016/j.jag.2015.10.002.

760   Martínez-Fernández, J., and A. Ceballos (2005), Mean soil moisture estimation using temporal stability

761   analysis, *J. Hydrol.*, *312*(1–4), 28–38, doi:10.1016/j.jhydrol.2005.02.007.

762   McColl, K. a., J. Vogelzang, A. G. Konings, D. Entekhabi, M. Piles, and A. Stoffelen (2014), Extended

763   triple collocation: Estimating errors and correlation coefficients with respect to an unknown target,

764   *Geophys. Res. Lett.*, *41*(17), 6229–6236, doi:10.1002/2014GL061322.

765 Merlin, O., C. Rüdiger, A. Al Bitar, P. Richaume, J. P. Walker, and Y. H. Kerr (2012), Disaggregation

766   of SMOS Soil Moisture in Southeastern Australia, *IEEE Trans. Geosci. Remote Sens.*, *50*(5), 1556–

767   1571, doi:10.1109/TGRS.2011.2175000.

768 Merlin, O., M. J. Escorihuela, M. A. Mayoral, O. Hagolle, A. Al Bitar, and Y. H. Kerr (2013), Self-

769   calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3 km

770   and 100 m resolution in Catalunya, Spain, *Remote Sens. Environ.*, *130*(2013), 25–38.

771 Miralles, D. G., W. T. Crow, and M. H. Cosh (2010), Estimating Spatial Sampling Errors in Coarse-

772   Scale Soil Moisture Estimates Derived from Point-Scale Observations, *J. Hydrometeorol.*, *11*,

773   1423–1429, doi:10.1175/2010JHM1285.1.

774 Mittelbach, H., and S. I. Seneviratne (2012), A new perspective on the spatio-temporal variability of soil

775   moisture: Temporal dynamics versus time-invariant contributions, *Hydrol. Earth Syst. Sci.*, *16*(7),

776   2169–2179, doi:10.5194/hess-16-2169-2012.

777 Mladenova, I. E., V. Lakshmi, T. J. Jackson, J. P. Walker, O. Merlin, and R. A. M. de Jeu (2011),

778   Validation of AMSR-E soil moisture using L-band airborne radiometer data from National

779   Airborne Field Experiment 2006, *Remote Sens. Environ.*, *115*(8), 2096–2103,

780   doi:10.1016/j.rse.2011.04.011.

781 Molero, B., O. Merlin, Y. Malbéteau, A. Al Bitar, F. Cabot, V. G. Stefan, Y. H. Kerr, S. Bacon, M. H.

782   Cosh, R. Bindlish, and T. J. Jackson (2016), SMOS disaggregated soil moisture product at 1 km

783   resolution: Processor overview and first validation results, *Remote Sens. Environ.*, *180*,

784   doi:10.1016/j.rse.2016.02.045.

785   Nicolai-Shaw, N., M. Hirschi, H. Mittelbach, and S. I. Seneviratne (2015), Spatial representativeness of

786   soil moisture using in situ, remote sensing, and land reanalysis data, *J. Geophys. Res. Atmos.*,

787   *120*(19), 9955–9964, doi:10.1002/2015JD023305.

788   Orlowsky, B., and S. I. Seneviratne (2014), On the spatial representativeness of temporal dynamics at

789   European weather stations, *Int. J. Climatol.*, *34*(10), 3154–3160, doi:10.1002/joc.3903.

790   Owe, M., R. A. M. de Jeu, and T. R. H. Holmes (2008), Multisensor historical climatology of satellite-

791   derived global land surface moisture, *J. Geophys. Res. Earth Surf.*, *113*(1), 1–17,

792   doi:10.1029/2007JF000769.

793   Pan, F., C. D. Peters-Lidard, and M. J. Sale (2003), An analytical method for predicting surface soil

794   moisture from rainfall observations, *Water Resour. Res.*, *39*(11), doi:10.1029/2003WR002142.

795   Parinussa, R. M., T. R. H. Holmes, and R. A. M. de Jeu (2012), Soil moisture retrievals from the

796   WindSat spaceborne polarimetric microwave radiometer, *IEEE Trans. Geosci. Remote Sens.*, *50*(7),

797   2683–2694, doi:10.1109/TGRS.2011.2174643.

798   Percival, D. B., and A. T. Walden (2000), *Wavelet methods for time series analysis*, Cambridge

799   University Press.

800   Rodriguez- Iturbe, I. (2000), Ecohydrology: A hydrologic perspective of climate-soil-vegetation

801   dynamics, *Water Resour. Res.*, *36*(1), 3–9, doi:10.1029/1999WR900210.

802   Smith, A., J. P. Walker, A. W. Western, R. I. Young, K. Ellett, R. C. Pipunic, R. B. Grayson, L.

803   Siriwardena, F. H. S. Chiew, and H. Richter (2012), The Murrumbidgee soil moisture monitoring

804   network data set, *Water Resour. Res.*, *48*(7), W07701, doi:10.1029/2012WR011976.

805 Stoffelen, A. (1998), Toward the true near-surface wind speed: Error modeling and calibration using

806   triple collocation, *J. Geophys. Res.*, *103*(C4), 7755, doi:10.1029/97JC03180.

807 Su, C.-H., and D. Ryu (2015), Multi-scale analysis of bias correction of soil moisture, *Hydrol. Earth*

808   *Syst. Sci.*, *19*(1), 17–31, doi:10.5194/hess-19-17-2015.

809 Su, C.-H., J. Zhang, A. Gruber, R. M. Parinussa, D. Ryu, W. T. Crow, and W. Wagner (2016), Error

810   decomposition of nine passive and active microwave satellite soil moisture data sets over Australia,

811   *Remote Sens. Environ.*, *182*, 128–140, doi:10.1016/j.rse.2016.05.008.

812 Vachaud, G., A. Passerat De Silans, P. Balabanis, and M. Vauclin (1985), Temporal Stability of

813   Spatially Measured Soil Water Probability Density Function, *Soil Sci. Soc. Am. J.*, *49*(4), 822,

814   doi:10.2136/sssaj1985.03615995004900040006x.

815 Vereecken, H., J. a. Huisman, Y. Pachepsky, C. Montzka, J. van der Kruk, H. R. Bogena, L.

816   Weihermüller, M. Herbst, G. Martinez, and J. Vanderborght (2014), On the spatio-temporal

817   dynamics of soil moisture at the field scale, *J. Hydrol.*, *516*(July 2015), 76–96,

818   doi:10.1016/j.jhydrol.2013.11.061.

819 Vinnikov, K. Y., A. Robock, N. A. Speranskaya, and C. A. Schlosser (1996), Scales of temporal and

820   spatial variability of midlatitude soil moisture, *J. Geophys. Res. Atmos.*, *101*(D3), 7163–7174,

821   doi:10.1029/95JD02753.

822 Vogelzang, J., and A. Stoffelen (2012), *Triple collocation, EUTMETSAT report*.

823 Wagner, W., V. Naeimi, K. Scipal, R. A. M. de Jeu, and J. Martínez-Fernández (2007), Soil moisture

824   from operational meteorological satellites, *Hydrogeol. J.*, *15*(1), 121–131, doi:10.1007/s10040-006-

825   0104-6.

826   Wang, G., D. Garcia, Y. Y. Liu, R. A. M. de Jeu, and a. Johannes Dolman (2012), A three-dimensional

827   gap filling method for large geophysical datasets: Application to global satellite soil moisture

828   observations, *Environ. Model. Softw.*, *30*, 139–142, doi:10.1016/j.envsoft.2011.10.015.

829   Wang, J., Y. Ge, G. B. M. Heuvelink, and C. Zhou (2015), Upscaling In Situ Soil Moisture Observations

830   To Pixel Averages With Spatio-Temporal Geostatistics, *Remote Sens.*, *7*(9), 11372–11388,

831   doi:10.3390/rs70911372.

832   Webster, R., and M. A. Oliver (1992), Sample adequately to estimate variograms of soil properties, *J.*

833   *Soil Sci.*, *43*(1), 177–192, doi:10.1111/j.1365-2389.1992.tb00128.x.

834   Western, A. W., G. Blöschl, and R. B. Grayson (1998), Geostatistical characterisation of soil moisture

835   patterns in the Tarrawarra catchment, *J. Hydrol.*, *205*(1–2), 20–37, doi:10.1016/S0022-

836   1694(97)00142-X.

837   Western, A. W., S. L. Zhou, R. B. Grayson, T. A. McMahon, G. Blöschl, and D. J. Wilson (2004),

838   Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial

839   hydrological processes, *J. Hydrol.*, *286*(1–4), 113–134, doi:10.1016/j.jhydrol.2003.09.014.

840   Wigneron, J.-P., Y. H. Kerr, P. Waldteufel, K. Saleh, M. J. Escorihuela, P. Richaume, P. Ferrazzoli, P.

841   de Rosnay, R. Gurney, J.-C. Calvet, J. P. Grant, M. Guglielmetti, B. K. Hornbuckle, C. Mätzler, T.

842   Pellarin, and M. Schwank (2007), L-band Microwave Emission of the Biosphere (L-MEB) Model:

843   Description and calibration against experimental data sets over crop fields, *Remote Sens. Environ.*,

844   *107*, 639–655, doi:10.1016/j.rse.2006.10.014.

845   Yee, M. S., J. P. Walker, A. Monerris, C. Rüdiger, and T. J. Jackson (2016), On the identification of

846   representative in situ soil moisture monitoring stations for the validation of SMAP soil moisture

847   products in Australia, *J. Hydrol.*, *537*, 367–381, doi:10.1016/j.jhydrol.2016.03.060.

848   Yilmaz, M. T., and W. T. Crow (2014), Evaluation of Assumptions in Soil Moisture Triple Collocation

849   Analysis, *Am. Meteorol. Soc.*, 1293–1302, doi:10.1175/JHM-D-13-0158.1.

850   Zwieback, S., K. Scipal, W. A. Dorigo, and W. Wagner (2012), Structural and statistical properties of

851   the collocation technique for error characterization, *Nonlinear Process. Geophys.*, *19*(1), 69–80,

852   doi:10.5194/npg-19-69-2012.

853

854

855 **Tables**

856 Table 1 – Wavelet scales for sampling period $T_s = 0.5$ days

Time scale j	Time scale (days) $\tau_j = 2^{j-1}\cdot T_s$
1	0.5
2	1
3	2
4	4
5	8
6	16
7	32
8	64
9	128

857

858 Table 2 - Values assigned to the Pan's model parameters for the generation on synthetic SM time series

Parameter	Value	Differences with Pan et al. [2003]
γ	0.40	Adjusted to control the effect of p ($\gamma = 1$ produced SM > 1 m^3/m^3)
SM_{min} (m^3/m^3)	0.025	
SM_{max} (m^3/m^3)	0.5	
η (m/yr)	$\eta(t) = \max\{0.2 \cdot Ks + 0.4 \cdot LAI(t), 0.5\}$	Equation changed to match observations
Z (m)	$Z = Z_{max} - r_{mod} \cdot (Ks - Ks_{min})$ with $r_{mod} = \frac{Z_{max}-Z_{min}}{Ks_{max}-Ks_{min}} = \frac{0.09-0.025}{5-0.05}$	Equation changed to match observations. A texture-depending Z allows a wider range of decay rates. Z is reduced as soil becomes sandier (smaller depth provokes faster changes)
Ks (cm/hr)	Sand: Ks = 5, Loam: Ks = 1.3	Source: FAO htp://ftp.fao.org/fi/cdrom/fao_training/FAO_ Training/General/x6706e/x6706e09.htm

859

860

861 Table 3 - Characteristics of the modeled SM series of Little Washita

Series	Variables		
	Precipitation	**Texture**	**LAI**
ref-a/b	$p_{ref}(t)$	loam	LAI_{ref-a} $LAI_{ref-b} = LAI_{ref-a}/4$
1-a/b	*Different amplitudes* $p(t) = p_{ref}(t) + N(0, \sigma_{p_{ref}}/4)$	loam	$LAI_{ref-a/b}$
2-a/b	*Different amplitudes and times: 10% of the events shifted +0.5 day and 10 %, -0.5 day* $p(t) = p_{ref-SHIFT}(t) + N(0, \sigma_{p_{ref}}/4)$	loam	$LAI_{ref-a/b}$
3-a/b	$p_{ref}(t)$	*sand*	$LAI_{ref-a/b}$
4-a/b	$p_{ref}(t)$	loam	*One-month shift* $LAI_{a/b}(t) = LAI_{ref-a/b}(t-30)$

862

863

864

865 **Figures**

866
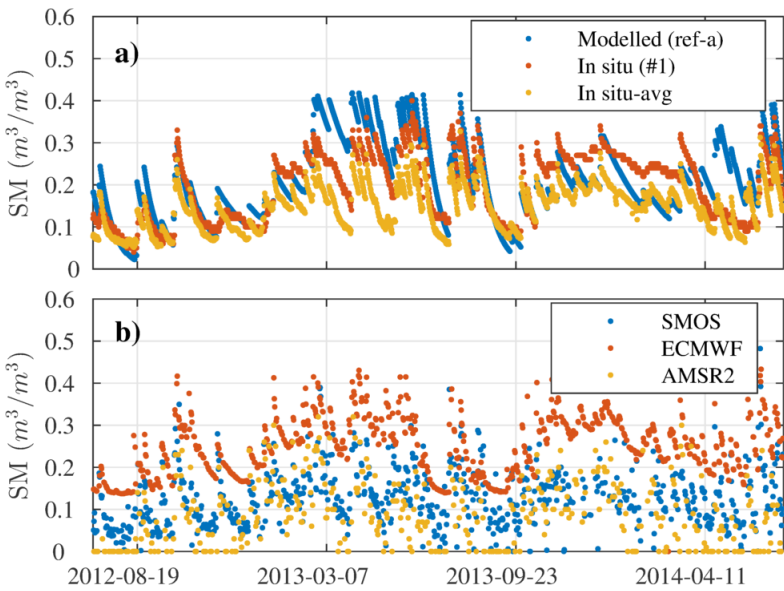
867 Figure 1 – Time series of the SM datasets used in the Little Washita region, before gap-filling. Only one of the
868 time series of the modeled dataset (ref-a) and two of the *in situ* dataset (station #1 and the spatial average) are
869 included.

870

871
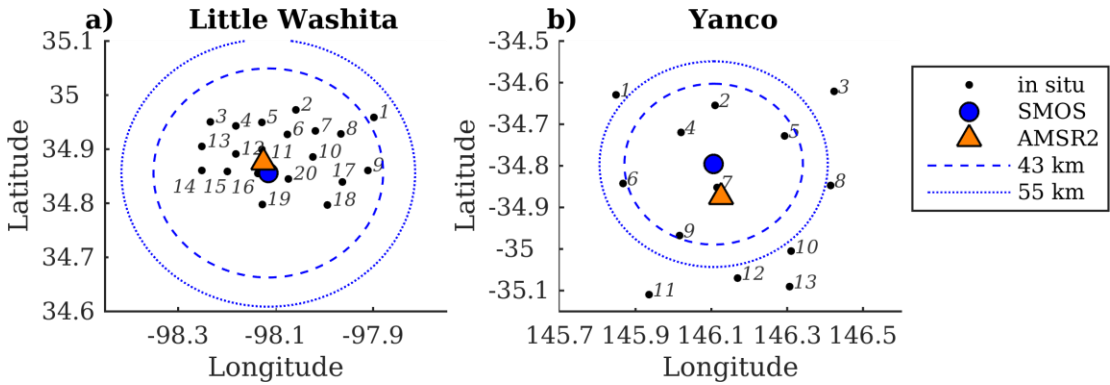
872 Figure 2 – Location of the *in situ* stations and the SMOS and AMSR2 grid nodes in each of the validation areas.
873 The circles represent two typical SMOS antenna footprints sizes considered in the retrieval algorithms: the
874 average one of 43 km and the maximum one of 55 km.

875

876
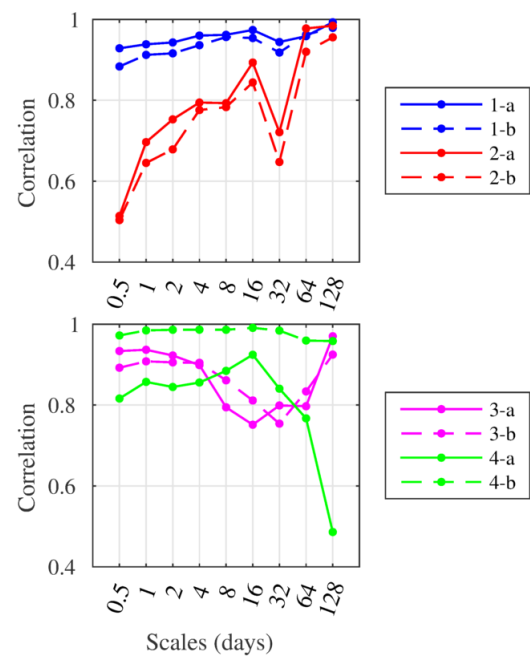
877 Figure 3 - Correlations between the detail series of different sample series and their respective reference series (a
878 or b), as a function of time scales.

879
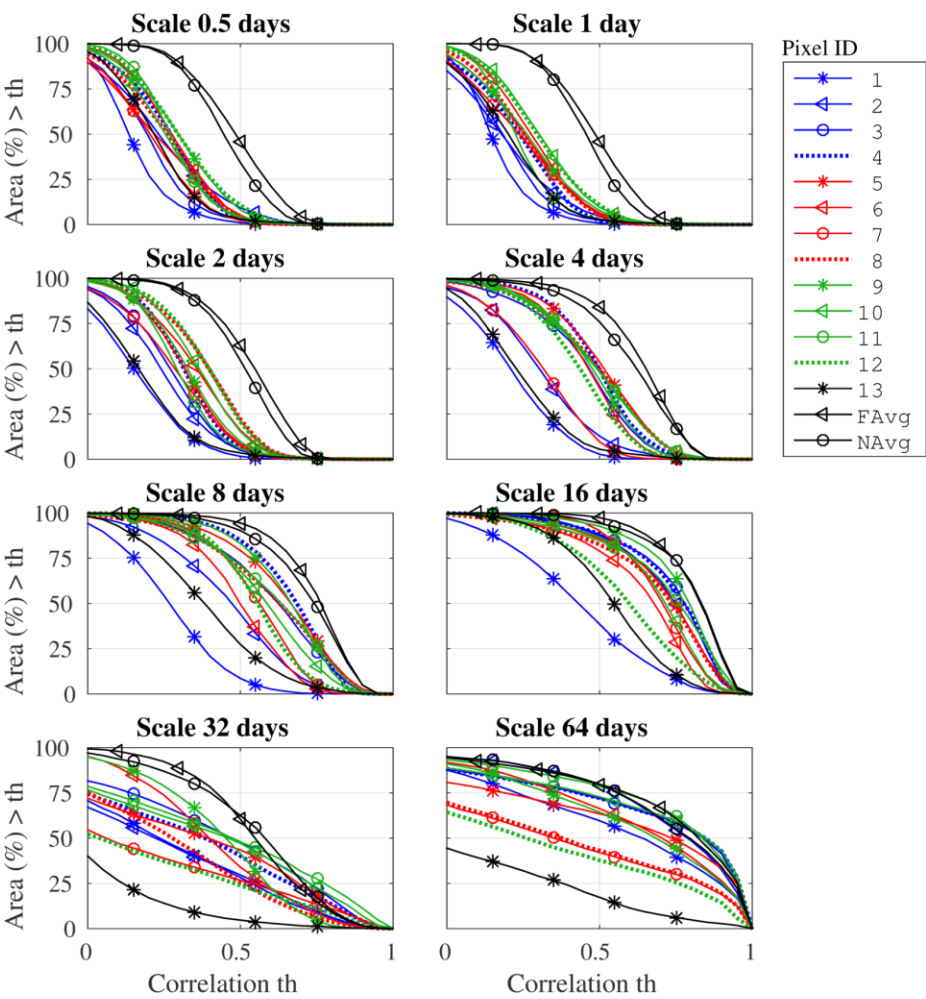
Figure 4 – CArea scores: size of the representativeness area in percentage of the total area, for a set of different locations (pixels) and the field and network average series (FAvg, NAvg). The dataset is the *in situ*-DIS Yanco dataset, for the 2014/09 - 2015/03 6-month period.
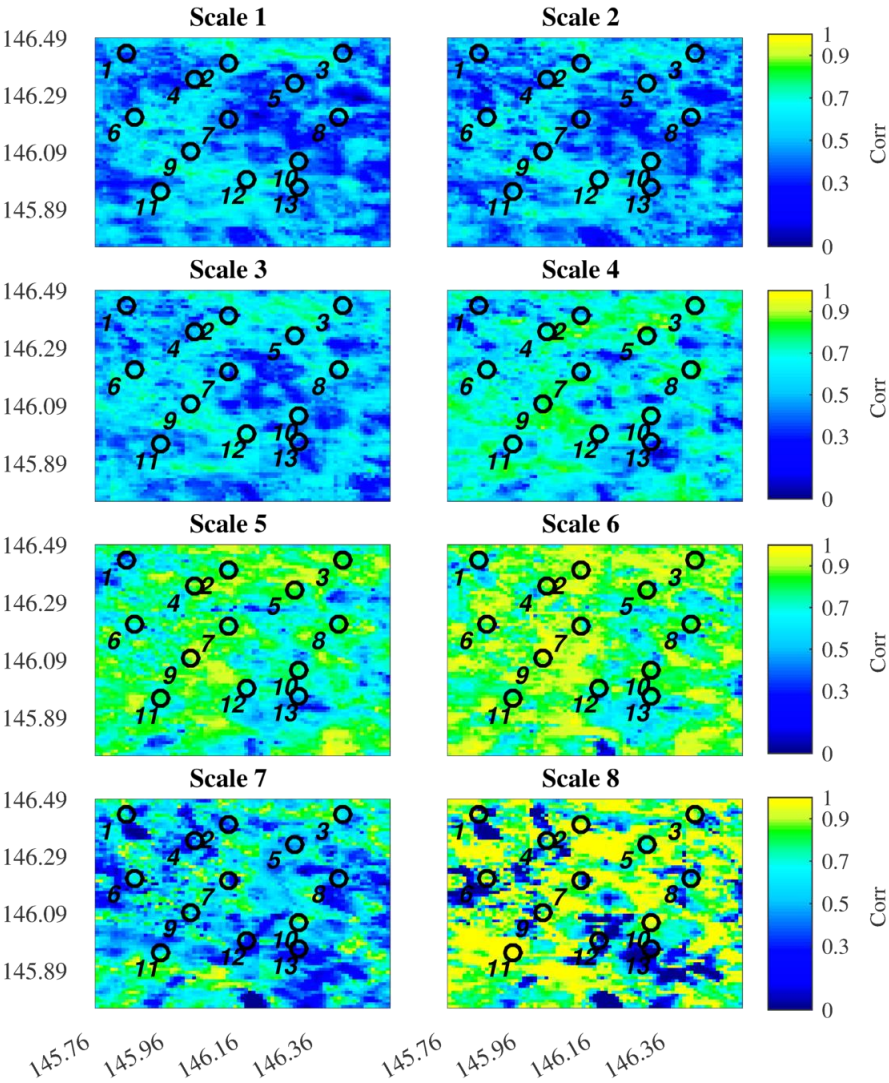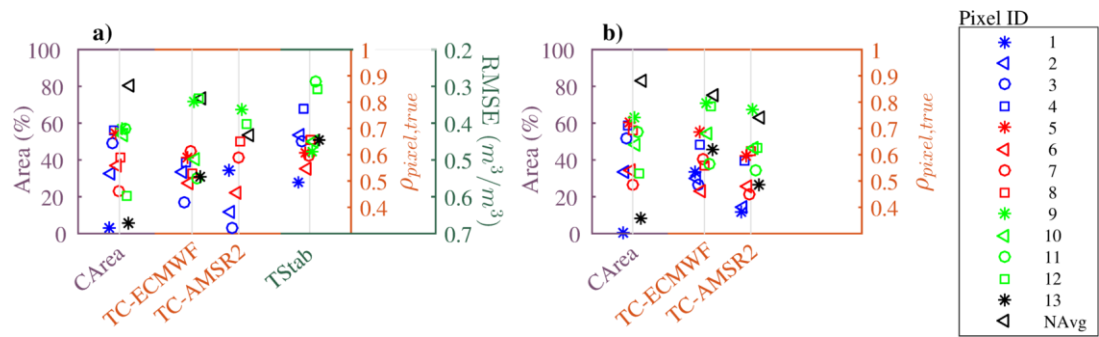
885

886　　Figure 5 - Maps of temporal correlation between each pixel time series and the field-average time series of the *in*
887　*situ*-DIS dataset. Values are calculated on detail series. Color code is bounded between 0 and 1, although negative
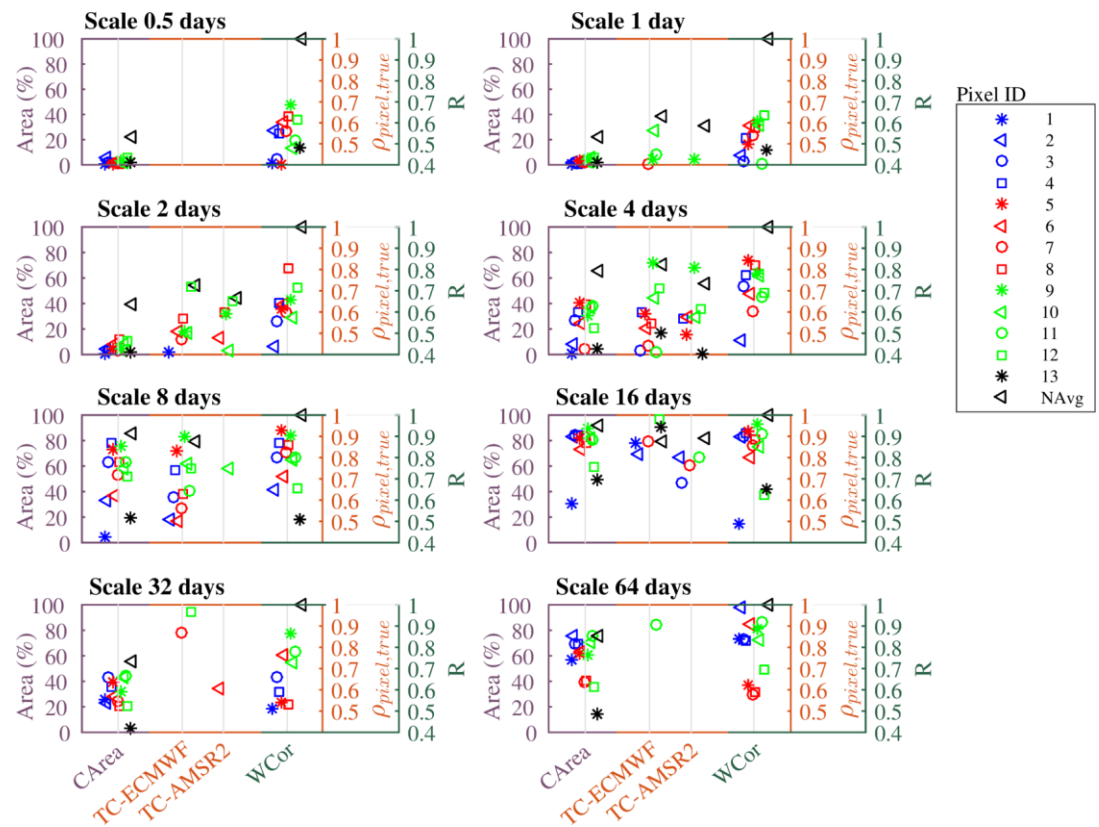888　　　　　　　　　　　　　　　　　　　　　correlation values exist.

889

890

891

Figure 6 - Representativeness values (vertical axis) from different methods (horizontal axis) for different pixels of the *in situ*-DIS Yanco dataset. The methods are applied to a) full series and to b) detrended series (components > 32 days are removed). The CArea correlation threshold is 0.55.
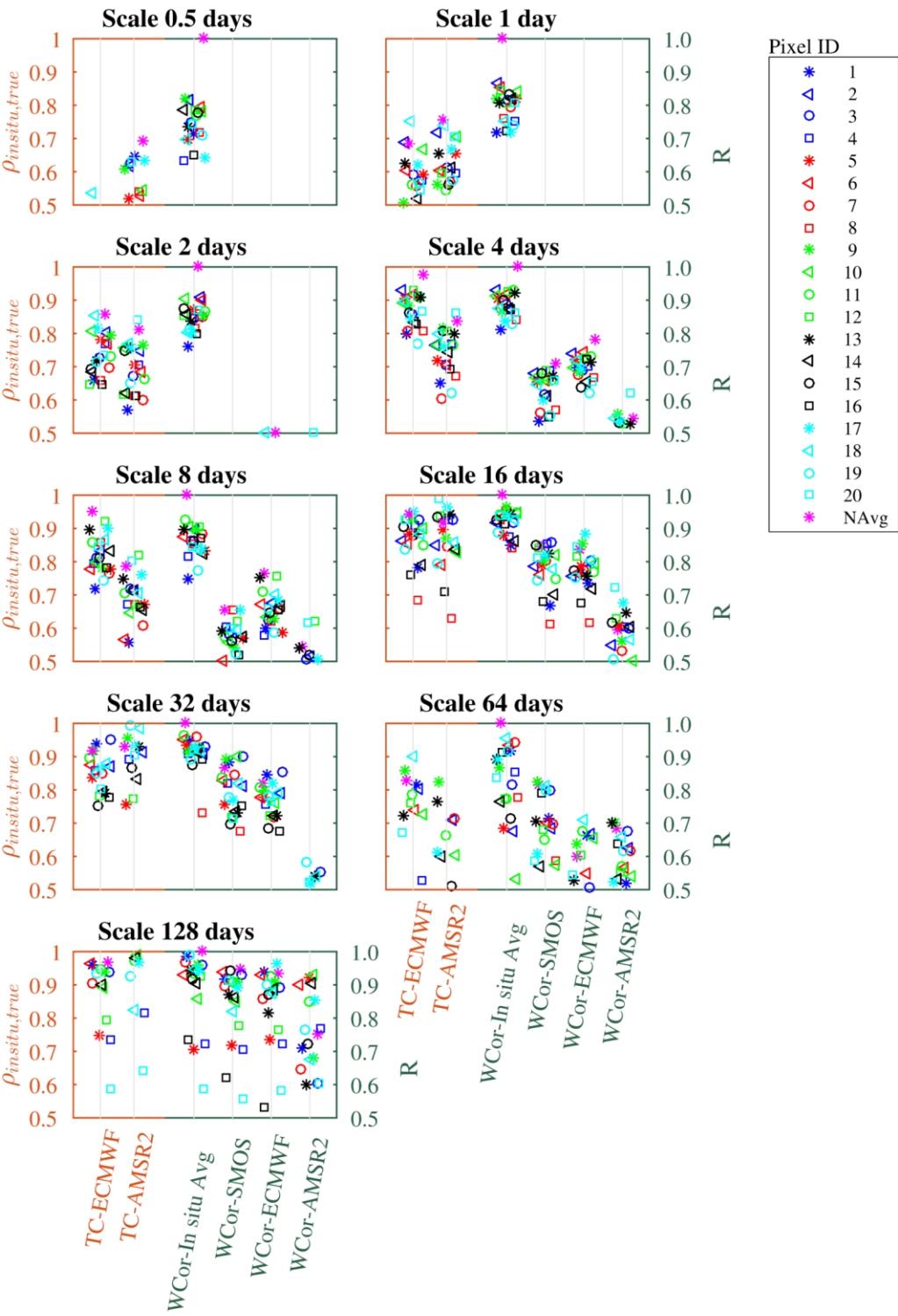
895

896

Figure 7 – Spatial representativeness values from CArea, TC and WCor methods for different pixels of the *in situ*-DIS Yanco dataset, per time scale. The CArea correlation threshold is 0.55.
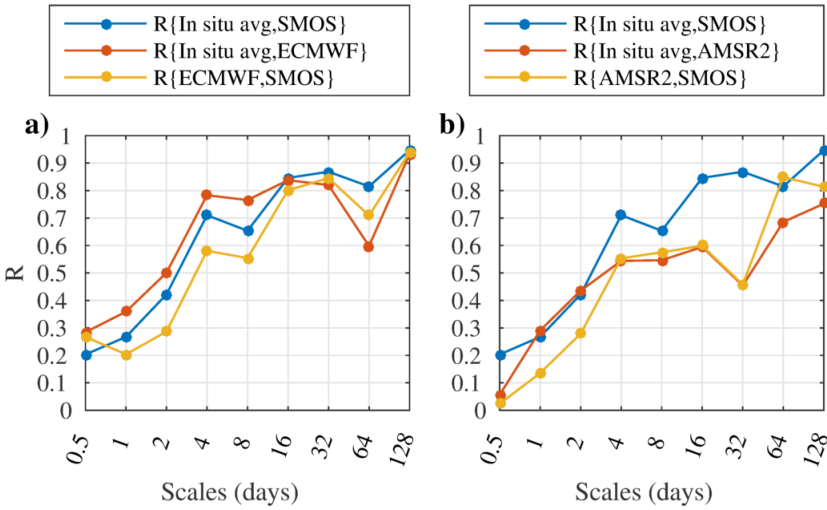
899

900

Figure 8 – Spatial representativeness of Little Washita stations per time scale, evaluated with different TC and WCor methods for the 2012/07 - 2014/07 2-year period

901
902

903

904

Figure 9 - Correlation between decomposed series (detail series) of the Little-Washita datasets: (a) the TC ECMWF-based triplet and (b) the TC AMSR2-based triplet. For clarity, only the *in situ* average series is present as *in situ* dataset.